DOCUMENT RESUME

ED 437 950 IR 057 579

TITLE	Digital Library Federation: Annual Report, 1998-1999.
INSTITUTION	Council on Library and Information Resources, Washington, DC. Digital Library Federation.
PUB DATE	1999-00-00
NOTE	14p.
PUB TYPE	Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Libraries; Annual Reports; *Electronic Libraries; Higher Education; Internet; *Library Administration; Library Cooperation; Library Development; *Library Materials; Library Services; Workshops

ABSTRACT

This annual report of the Digital Library Federation describes projects initiated in response to the following program priorities: (1) focus attention on libraries of materials that are born digital, including a social science data workshop and guides to quality in visual resource imaging; (2) integrate digital materials into the fabric of academic life, including academic image exchange, authorization systems, reference linking, and distributed finding aids; (3) stimulate the development of core digital library infrastructures, including a workshop on technical metadata elements for image files, international meeting on archival authority control, metadata working group, technical architecture committee, and I2-DSI (Internet2--Distributed Storage Infrastructure); and (4) develop the organizational support needed for managing digital libraries effectively, including institutional policies and practices for the long-term maintenance of digital materials and management forums. Outreach initiatives (e.g., publications and site visits) and organizational issues (e.g., executive subcommittee, new members, alliances, and staff) are also described. (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 437 950

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

B.H. Leneyr

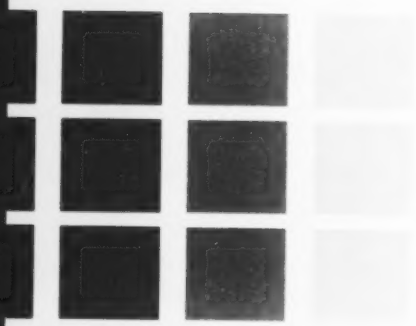
TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.



DIGITAL LIBRARY FEDERATION

ANNUAL REPORT 1998-1999

IR057579

BEST COPY AVAILABLE

DIGITAL LIBRARY FEDERATION

PARTNERS

California Digital Library
Carnegie-Mellon University
Columbia University*
Cornell University*
Council on Library and Information Resources*
Emory University*
Harvard University*
Indiana University
Library of Congress*
New York Public Library*
North Carolina State University
Pennsylvania State University*
Princeton University*
Stanford University*
University of California, Berkeley*
University of Chicago
University of Michigan*
University of Minnesota
University of Pennsylvania
University of Southern California*
University of Tennessee*
University of Texas at Austin
Yale University*

ALLIES

Coalition for Networked Information
National Archives and Records Administration*
Online Computer Library Center, Inc. (OCLC)
Research Libraries Group

*Founding partner

The Digital Library Federation is administered by the Council on Library and Information Resources (CLIR), a not-for-profit organization dedicated to supporting the scholarly communication system upon which knowledge creation depends. CLIR aims to identify the critical issues affecting the welfare and prospects of libraries and archives and the constituencies they serve, convene individuals and organizations in the best position to engage these issues and respond to them, and encourage institutions to work collaboratively to achieve and manage change. CLIR's projects and activities aim to ensure that information resources needed by scholars, students, and the public, are available today and for future generations.

BEST COPY AVAILABLE

INTRODUCTION

The Council on Library and Information Resources (CLIR) is deeply interested in fostering the development of digital libraries as a resource for research and learning. We are particularly interested in helping policy makers, funding organizations, and academic leaders understand what social and institutional investments in digital libraries are necessary to organize, maintain, and provide access to the growing body of digital materials for scholarly purposes. The development of digital libraries is proving to be very complex, requiring much intensive and detailed work. At present, we seek to help accomplish this work and advance our interest in digital libraries primarily by serving as administrative home to the Digital Library Federation (DLF).

The Digital Library Federation was formed in 1995 as a consortium of 15 research institutions with a common goal to create a system of independent, distributed repositories for digital works. The founding members of the federation aspired to develop such distributed repositories into a globally accessible virtual library for research and education, one that provides good local service while ensuring that the value of the distributed works overall is more than the sum of its parts. The members of the federation committed themselves to a shared investment in developing the means to "federate" libraries of digital works.

An important first step for the DLF was to agree on a definition of digital libraries. Its members agreed on a broad interpretation:

> Digital libraries are organizations that provide the resources, including specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.

At the end of the fiscal year, DLF partners numbered 23. In addition, there are four DLF allies: the Coalition for Networked Information (CNI), the National Archives and Records Agency (NARA), OCLC, Inc., and the Research Libraries Group (RLG). During his first year of work as director of the Digital Library Federation, Donald Waters visited DLF partners and allies to learn about their organization and their existing and planned projects. The visits convinced him that the success of digital library efforts will be determined by the extent to which they address or help achieve the following goals:

- organizing, providing access to, and preserving knowledge that is born digital
- leveraging digital library facilities for managing intellectual works in support of efforts to redesign the scholarly communication process
- providing an accessible and durable knowledge base that helps improve the quality and lower the costs of education
- extending the reach of research and higher education to new segments of the citizenry

With these goals in mind, the Digital Library Federation agreed on four broad program priorities:

- focus attention on libraries of materials that are born digital
- integrate digital materials into the fabric of academic life
- stimulate the development of a core digital library infrastructure
- develop the organizational support needed for managing digital libraries effectively

The DLF initiated the following projects in response to these program priorities.

Focus on Materials Born Digital

Social Science Data Workshop

DLF organized and conducted a workshop on social science data at Princeton University in January 1999. Social science data managers from DLF institutions joined a variety of experts to examine the state of the art in three areas: the discovery and retrieval of databases, the evaluation and interpretation of alternative data sources,

and data extraction for analysis and presentation. Speakers included prominent faculty in the field, such as Gary King, of Harvard; Richard Rockwell, of the Inter-university Consortium for Political Science Research (ICPSR); and Daniel Greenstein, of the Arts and Humanities Data Service in the United Kingdom. Participants identified a set of activities that the DLF can undertake to advance the state of the art in these three areas with the goal of improving the use of social science databases in the undergraduate curriculum. The workshop report[1] outlines an agenda for action that includes the following:

- There are not enough skilled staff, nor is the infrastructure sufficient, to create and disseminate SGML- and XML-encoded documentation on a large scale. To address the problem, library and other institutional leaders must be informed about quantitative service needs, and encouraged to do strategic planning to meet the needs of research and teaching within and across disciplines.

- Independent, uncoordinated, and duplicative work is under way on most campuses that are developing mechanisms that enable faculty and students to use key data sets. These access mechanisms must be designed to allow scientific replication of data analysis. One solution would be to devise a master plan for systematically dividing the labor of developing access tools among a variety of campuses. An alternative would be for institutions to contribute work in a common but distributed framework of data repositories and tools.

- The development and deployment of a new, SGML-based standard for documenting data sets in codebooks will require campus investments, the conversion of existing codebooks to digital form, and research to understand how online codebooks will relate to other discovery and access tools for numeric data.

- Strategies are needed to address the dual challenge of preserving digital data: maintaining tools to read the digital data files, while also preserving the codebooks needed to interpret the data output.

- Tools and other facilities must be developed to help users understand and deal with comparability across related data sets.

- More attention should be paid to the cataloging of data sets.

- Consortia should be developed—or existing consortia used—to negotiate the purchase and licensing of critical and expensive data sets.

---

[1] The report is available at **www.clir.org/diglib/ssdaresults.htm**.

Guides to Quality in Visual Resource Imaging

The DLF, CLIR, and RLG have created an editorial board of experts to review the state of the art in visual resource imaging and to identify technologies and practices that can be documented and recommended to the community. The board decided to focus on documenting the science of imaging; that is, some of the objective measures of image qualities, such as color, tone, and resolution, and how they can be controlled in various aspects of an imaging process. It identified five areas in which to address these issues: setting up an imaging project, selecting a scanner, creating a scanning system, producing a digital master, and generating digital derivatives. Board members created detailed outlines for guides in these areas and suggested authors, whom DLF and CLIR commissioned to write the guides. The guides will be published on the RLG Web site in late 1999.

Integrate Digital Materials into the Fabric of Academic Life

Academic Image Exchange

In January 1999, DLF convened a meeting to explore ways of using digital libraries to enhance the quality of art history teaching and research in the nation's colleges and universities. Art history faculty, visual resources librarians, and representatives of the College Art Association attended. As a result of the meeting, participants are proceeding with the development of a prototype mechanism, called an image exchange. The facility would enable scholars to share images, to which they own the rights, of works referenced in the major art history textbooks. Since March, project participants have opened discussions of the development of the image exchange through formal presentations at the CNI Spring Meeting and a regional meeting of the Visual Resources Association, and through informal presentations to other interested organizations and art history publishers. In addition, DLF contracted with Robert Baron, an independent art historian and consultant with expertise in information technology and intellectual property matters, to serve as project manager. In May, the image exchange organizing group met at Carnegie Mellon University to design and begin constructing the prototype system. With the help of CMU computer scientists, the group settled on a comprehensive set of features for the exchange. The group is now developing these features, including a concordance of images in key art and architectural history textbooks, for the prototype. For additional details see **www.clir.org/diglib/artxdescription.htm**.

Authorization Systems

Digital certificates offer a secure means of authorizing access to a range of campus systems and resources and are becoming part of campus technology infrastructure. Under DLF auspices, the University of California, Columbia University, JSTOR, and OCLC are developing a protocol that

will enable an information resource provider to verify that a user bearing a digital certificate has authority from a home institution to use a requested resource. The prototype system being developed combines the use of X.509 digital certificates for authentication with a directory service providing authorization to licensed resources based on user attributes. The first phase of work on the project was presented at the CNI Spring Meeting in Washington, D.C., in April 1999. Following the meeting, project participants met and outlined a work plan for a second phase of development.

The group focused on the need to recruit additional participants for further testing and development of the protocol. They devised a strategy to engage additional participants by aiming at campus technology organizations as well as at libraries. The DLF has worked with the Corporation for Research and Educational Networking (CREN) to develop a paper of frequently asked questions and answers about the use of digital certificates that is aimed at senior campus administrators. In addition, project participants presented the project to the IT directors attending the spring meeting of the Common Solutions Group. In early June, project members attended a meeting of the Committee on Institutional Cooperation's Research Projects Group, which is charged with making recommendations to the chief information officers of that consortium. As a result of these efforts, several additional institutions have been identified as potential participants in the project.

To engage more vendors in the project, Columbia and the California Digital Library are identifying those with whom they currently have agreements for access to licensed resources. Administrators of the Joint Information Systems Committee in the U.K., who are watching this project closely, have suggested several vendors who are working with them on authorization systems.

Also in the project's second phase, the project group is drafting a digital library authentication and authorization architecture statement that will be presented to the Internet Engineering Task Force later this year for consideration as a Request for Comment.

Reference Linking

In February 1999, the DLF joined the National Information Standards Organization (NISO), the Society of Scholarly Publishers (SSP), and the National Federation of Indexing and Abstracting Services (NFAIS) in sponsoring a workshop on linkage from citations to the electronic journal literature. The workshop brought together publishers, librarians, representatives from abstracting and indexing services, information aggregators, vendors of information services, and end users to build a common aware-

ness of a broad range of needs and to improve understanding of the strengths and limitations of current approaches. In addition, the workshop sought to identify and stimulate actions needed to improve the facilities for linking citations and digital objects in the digital environment.

As a result of the workshop, the sponsors created a small working group to explore the issues more fully. The working group sought to define the nature and scope of research and user consultation, and to identify other work necessary to foster the development of general systems of reference linking for marketplace testing. The report of the group was issued in May 1999 and is available at **www.lib.uchicago.edu/Annex/pcaplan/reflink.html**. The report served as the basis for a second invitational workshop held in June 1999. Among the results of the second workshop was a plan for additional research by computer scientists. The DLF led a call for development of solutions to the problem of selective resolution when there are multiple copies of a cited work. For more details about the second workshop on reference linking, see **www.niso.org/linkge2.html**.

Distributed Finding Aids

This project, under the direction of John-Price Wilkin at the University of Michigan, is exploring the means and costs of searching encoded finding aids that are distributed at different institutions. The final report is available at **www.clir.org/diglib/dfas.htm**.

Stimulate the Development of Core Digital Library Infrastructures

Workshop on Technical Metadata Elements for Image Files

In April 1999, NISO, CLIR, and RLG sponsored an invitational workshop to examine the technical information needed to manage and use digital still images that reproduce a variety of pictures, documents, and artifacts. Sixty individuals with diverse interests and perspectives on the problem of metadata information attended the meeting. DLF member institutions were well represented among the attendees, which included libraries, universities, museums and archives, as well as representatives from government, other digital library organizations, and the digital imaging vendor community. By the end of the meeting, the participants reached agreement on a variety of issues, including the following:

- a preliminary list of technical metadata elements
- the use of industry standard metrics, where they exist, for assessing images (for example, tone, color, International Color Consortium (ICC) profiles)

- the need to develop methods of pointing to external test charts
- a requirement for mechanisms enabling the metadata associated with an image to persist through various transformations

The draft report can be found at **www.niso.org/image.html**.

International Meeting on Archival Authority Control

In December 1998, the DLF sponsored a meeting of an international group of archivists to develop a standard encoding format for the recording and exchange of archival authority information. Their intention is that such a format would become part of the emerging archival information architecture that currently consists of formats for collective description (e.g., MARC-AMC) and for detailed finding aid descriptions, especially those using the Encoded Archival Description (EAD) format. The establishment of a standard for archival authority records is expected to have a major effect on the future direction of archival description and to resolve many long-standing questions about how best to record and present archival data. The meeting resulted in the development of a plan outlining the steps necessary to formulate an archive authority standard. Identified as being of particular importance in accomplishing adoption of such a standard was the need to incorporate recommendations into the International Standard Archival Authority Record for Corporate Bodies, Persons, and Families [ISAAR (CPF)]. The ISAAR (CPF) standard will be reviewed in 2001, providing an opportunity for incorporation of efforts by this group and its participants. Additional tasks identified in the project plan were the drafting of an ISAAR-compliant document-type definition (DTD), definition of a Z39.50 attribute set for ISAAR, and ensuring links to specialized authority files such as those for geographic information.

Metadata Working Group

The DLF convened an initial meeting of metadata experts in February 1999. The group considered how to best integrate, at descriptive and subject levels, the important Americana being created in digital form. The idea to develop an "academic Lycos" emerged in this meeting and was defined in a preliminary way as a project for subsequent work. The group also called for an overview of issues associated with the recording and use of structural metadata. And it began planning a series of meetings that would help familiarize practitioners at DLF institutions with metadata developments in the computer science labs associated with the NSF Digital Library Initiative and other projects.

Technical Architecture Committee

Following a tradition of meeting with individuals or institutions that have developed interesting approaches to digital library architecture, the

Technical Architecture Committee met with staff of OCLC in November 1998. The committee explored in detail OCLC's planned development of a Cooperative Online Resource Catalog (CORC) and recommended that DLF convene a metadata committee in part to explore possible DLF contributions to the CORC project. A member of JSTOR's technical development staff also attended the meeting and joined in a discussion of requirements that need to be met for publishers and libraries to begin using X.509 standard certificates as a means of authorizing users to gain access to licensed resources. In the course of the discussions, the committee formulated an authorization protocol to be developed under the auspices of DLF. Minutes of the November meeting are posted on the DLF Web site.

Internet2—Distributed Storage Infrastructure (I2-DSI)

The Innovative Computing Laboratory in the Department of Computer Science at the University of Tennessee is designing an architecture and a set of services to dramatically improve the delivery of large sets of data in a highly distributed environment. In theory, the Internet2 infrastructure is especially well suited to digital library applications. With partial support from the DLF, Micah Beck and Terry Moore, of the Innovative Computing Laboratory, and Bert Dempsey, of the School of Information and Library Science at the University of North Carolina at Chapel Hill, organized a workshop in March to explore how well the emerging infrastructure meets the needs of various applications, including those of digital libraries. Workshop participants considered several projects that could test the new infrastructure. Among the applications reviewed were a project for publications in the earth sciences being developed at Columbia University and a project to deliver online access to sound recordings being developed at Indiana University. A participant in the Academic Image Exchange project also made a presentation, and the project will be among the applications receiving support from the I2-DSI development team. More information is available at **http://dsi.internet2.edu/apps99.html**.

Develop the Organizational Support Needed for Managing Digital Libraries Effectively

Institutional Policies and Practices for the Long-term Maintenance of Digital Materials

The DLF and RLG have formed a task force to advance the understanding of digital preservation policy and practice. Few institutions have formal policies in place but many have made commitments to preserve digital materials and are adopting practices to support the handling and long-term maintenance of such materials. The goal of the task force is to identify and document these practices so that DLF and RLG can identify what kinds of technical, economic, and organi-

zational barriers institutions are facing, and what kinds of actions are needed to reduce those barriers.

The task group will gather and analyze existing digital preservation policies and practice descriptions for the following three classes of electronic materials:

- institutional records in digital form
- locally digitized materials
- electronic publications

Management Forums

In June, DLF staff made final preparations for the Digital Library Federation's first semi-annual Forum on Digital Library Practices, scheduled for July 1999. The agenda included presentations on authentication and authorization systems, digital repositories, finding aids, page image navigation systems, and naming systems.

Outreach Initiatives

Publications

In February 1999, the Digital Library Federation issued its first publication, *Enabling Access in Digital Libraries: A Report on a Workshop on Access Management*, edited by Caroline Arms with Judith Klavans and Donald Waters.

The second publication in the DLF series, *Preserving the Whole: A Two-track Approach to Rescuing Social Science Data and Metadata*, by Ann Green and JoAnn Dionne, appeared in June 1999.

Site Visits

Donald Waters visited the California Digital Library and the University of Chicago libraries. Rebecca Graham visited Carnegie Mellon libraries, the Yale University Library, California Digital Library, the libraries at the University of California at Berkeley, the Princeton University Library, the Library of Congress, and the Columbia University Library.

Organizational Issues

Executive Subcommittee

An executive subcommittee of the DLF Steering Committee was formed and began work in January with a dual purpose. First, it is charged to help plan the Steering Committee meetings. Second, it provides a ready source of advice for the DLF staff as needed between Steering Committee meetings on such matters as applications

from prospective new members, specific program plans, budgetary matters, positions that DLF should take when consulted by outside groups, and so on. The members of the executive subcommittee for 1999 are Paula Kaufman (chair), Richard Lucier, Elaine Sloan, Winston Tabb, and Sarah Thomas.

New Members

The University of Texas at Austin joined the DLF in January. The University of Virginia is preparing an application.

Alliances

CNI has joined in an alliance with DLF, on the same terms as RLG, OCLC, and NARA. That is, CNI and DLF will develop work jointly where their interests intersect, and Clifford Lynch will represent CNI at Steering Committee meetings as an ally "with voice but without vote."

Staff

Rebecca Graham joined the CLIR staff as DLF research associate in September 1998. Donald Waters, director of the DLF since October 1997, resigned on June 30, 1999, to accept the position of program officer for scholarly communication at The Andrew W. Mellon Foundation.

COUNCIL ON LIBRARY AND INFORMATION RESOURCES

1755 MASSACHUSETTS AVENUE, NW, SUITE 500, WASHINGTON, DC 20036-2124
Telephone: 202.939.4750 • Fax: 202.939.4765 • E-mail: info@clir.org • Web: www.clir.org